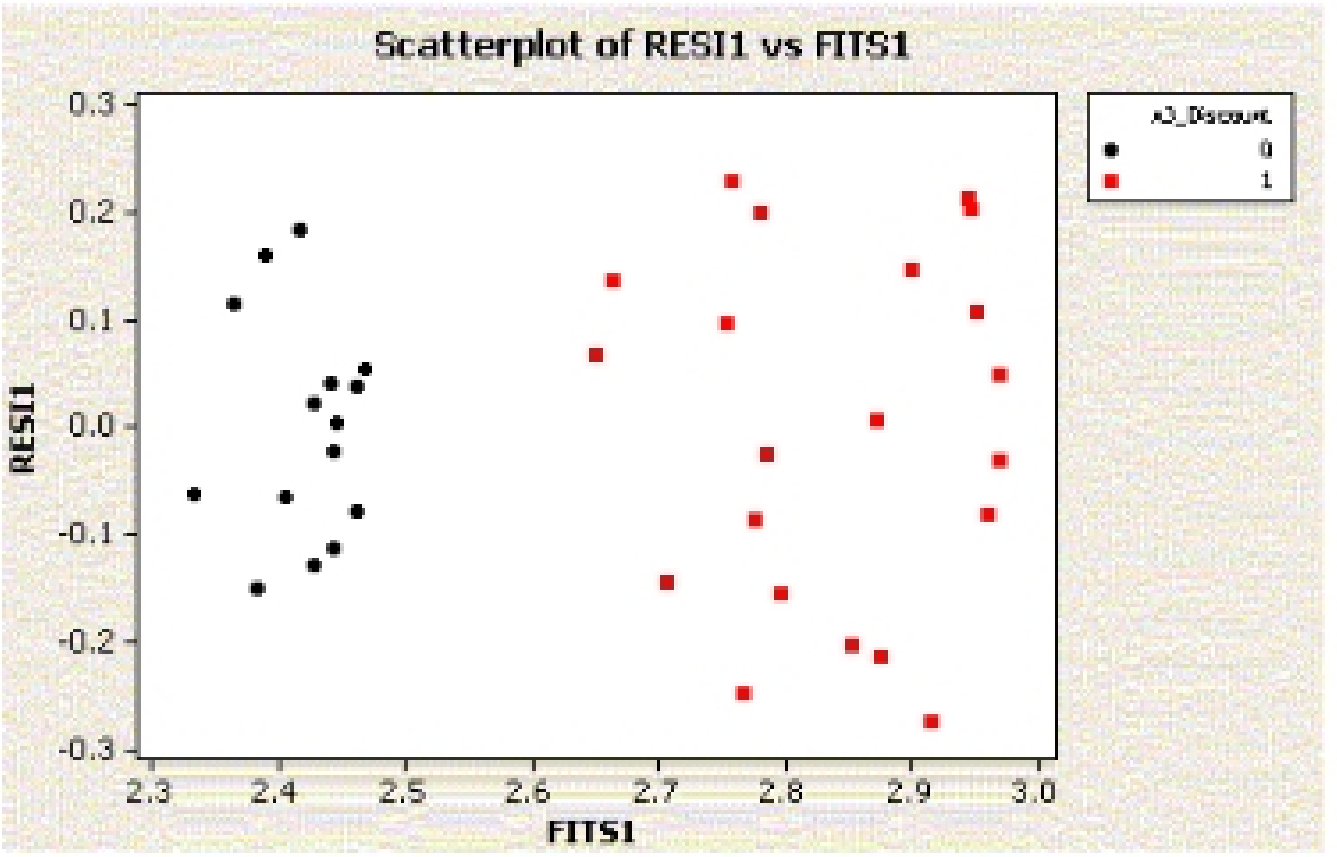
Stat 501 November 29 Nonconstant Variance and Weighted Least Squares

Example 1.
Y_Price = market share of product;
x1_price = price;
x3_Discount = 1 if discount promotion in effect and 0 otherwise;
x3x4=1 if both discount and package promotions in effect and 0 otherwise. N = 36 consecutive months.

Regression results including residual plot that indicates values of x3:

Predictor	Coef	SE Coef	T	P
Constant	3.1959	0.3562	8.97	0.000
x1_price	-0.3336	0.1523	-2.19	0.036
x3_Discount	0.30808	0.06412	4.80	0.000
x3x4	0.17623	0.06597	2.67	0.012

Source	DF	SS	MS	F	P
Regression	3	1.76282	0.58761	27.51	0.000
Residual Error	32	0.68343	0.02136		
Total	35	2.44626			



Interpretation: There's a pattern of nonconstant variance related to whether x3 = 1 or x3 = 0. Here are the residual variances for the two groups.

Variable	x3_Discount	N	N*	Variance
RESI1	0	15	0	0.0105
	1	21	0	0.0268

Weighting for a weighted least squares: For weights, we use $\frac{1}{\hat{\sigma}_i^2}$. In this case the two variances given just above estimated the σ_i^2. This leads to these weights:

Variable	x3_Discount	Weight
wts	0	95.238
	1	37.313

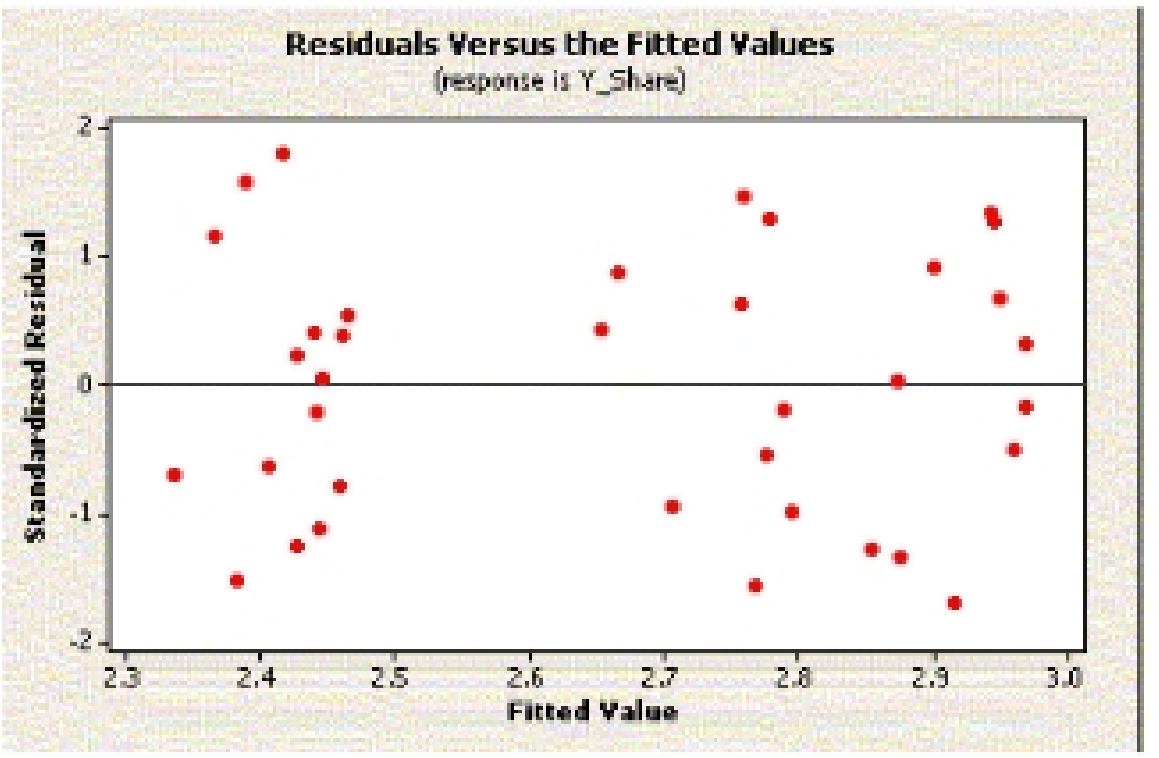
Weighted analysis using weights in wts

Predictor	Coef	SE Coef	T	P
Constant	3.1743	0.3567	8.90	0.000
x1_price	-0.3243	0.1529	-2.12	0.042
x3_Discount	0.30834	0.06577	4.69	0.000
x3x4	0.17585	0.07612	2.31	0.027

Source	DF	SS	MS	F	P
Regression	3	98.118	32.706	30.75	0.000
Residual Error	32	34.031	1.063		
Total	35	132.149			

IMPORTANT: Minitab's AOV will be in terms of weighted SS. Notice how different the SS values are from the SS values for the unweighted case.

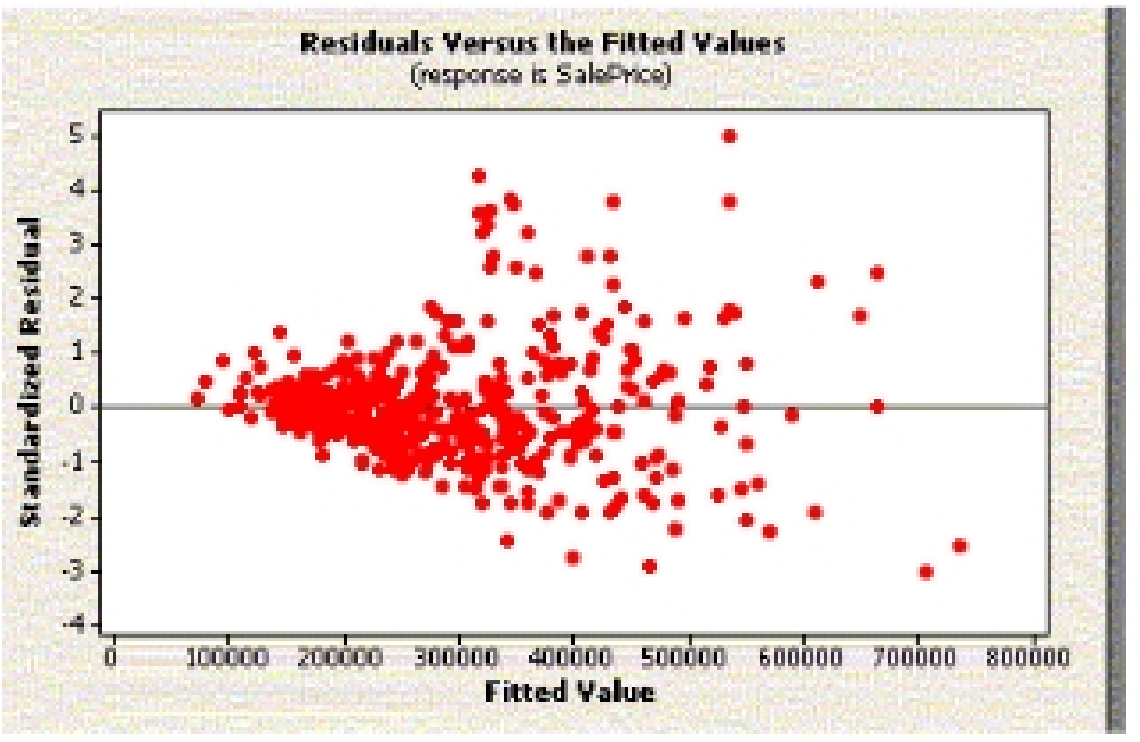
IMPORTANT: When plotting residuals after a weighted least squares, use standardized residuals.



Example 2. Y = Sale Price of Home, X1 = SqrFeet = square feet size of home, X2 = Lot Size = square feet lot size.

Regression results and residual plot:

Predictor	Coef	SE Coef	T	P
Constant	-102610	12531	-8.19	0.000
SqrFeet	156.228	4.825	32.38	0.000
Lotsize	1.1012	0.2940	3.75	0.000



Interpretation: There's a strong pattern of nonconstant variance related to the mean (Fit).

Two possibilities for weights:
1. Assume variance is proportional to mean, so use $w_i = 1/\text{fits}$.
2. Assume standard deviation is proportional to mean, so use $w_i = 1/(\text{fits})^2$

Results after using possibility 1 (residual polot on other side):

Predictor	Coef	SE Coef	T	P
Constant	-92713	10565	-8.78	0.000
SqrFeet	153.264	4.731	32.39	0.000
Lotsize	0.9699	0.2631	3.69	0.000